1-1-1993

Perceptual Bases of Visual Literacy

Paul Messaris
University of Pennsylvania, pmessaris@asc.upenn.edu

Recommended Citation (OVERRIDE)
Messaris, P. "Perceptual Bases of Visual Literacy." (1993). In R. A. Braden, J. C. Baca, & D. G. Beauchamp (Eds.), Art, Science and Visual Literacy: Selected Readings from the 24th Annual Conference of the International Visual Literacy Association (pp. 245-251). Blacksburg, VA: International Visual Literacy Association.

Perceptual Bases of Visual Literacy

by
Paul Messaris

Anyone with a scholarly interest in visual literacy is likely to have come across more than one story about people from pre-industrial societies who reportedly were unable to make any sense of their first encounters with photographs -- even when those photographs were of very familiar subjects (e.g., see Deregowski, 1980; Segall et al., 1966). Such accounts may be seen as evidence in favor of the proposition that the representational conventions of photographs and other kinds of pictures are largely or wholly arbitrary, despite the fact that viewers familiar with a particular pictorial style may think of it as being "natural" or "true to life." This view of pictures as an arbitrary system of signification was expressed as early as 1925 by the art historian Erwin Panofsky, in a celebrated essay on Renaissance perspective, but its most influential exponent has probably been the philosopher Nelson Goodman (1976).

Image-Reality Discrepancies

As Panofsky, Goodman, and other writers (e.g., Gombrich, 1960; Snyder, 1980; Wartofsky, 1984) have pointed out, many pictorial styles do indeed entail sharp discrepancies between image and reality. A minimal list of these discrepancies would encompass the following three broad areas:

(1) *Discrepancies in Color and Illumination*. For example, black-and-white photographs contain no information about the hue of colors, while unshaded outline drawings are uninformative about both color and illumination.

(2) *Discrepancies in Depth Cues*. Because pictures which attempt to represent three-dimensional scenes must do so on a flat, two-dimensional surface, all pictures lack some of the informational cues from which we infer the third dimension in real-world perception. (The nature of these cues will be described below.)

(3) *Discrepancies in Shape*. In many kinds of pictures, across a broad range of cultures and styles, the shapes of pictured objects are conventionally rendered very differently from the shapes of their real-world counterparts. Stick figures and most children's cartoons are good examples of this process.

In view of such discrepancies, it may seem intuitively reasonable to suppose that viewers who are not familiar with a particular pictorial style will be likely to have problems in interpreting pictures in that style. In other words, it may seem that we have here a good argument in support of the conclusion that pictorial conventions containing such discrepancies are indeed arbitrary, as Panofsky, Goodman, and their followers have claimed. As it happens, however, this conclusion has been contradicted repeatedly by a growing body of empirical evidence.

Over the past thirty years, more than a dozen studies have directly investigated how inexperienced viewers -- including children, people from societies without mass media, and even animal subjects -- respond to each of the three types of pictorial discrepancy listed above. Contrary to what one might have expected -- and contrary, also, to the outcomes of the anecdotal reports mentioned earlier -- these studies have found that inexperienced viewers are typically *not* troubled by the first and third types of discrepancy. In other words, they typically do *not* find it difficult to recognize the content of such things as black-and-white photographs, outline drawings, stick figures, and cartoons (provided the objects represented in these pictures are culturally familiar, of course). Only the perception of depth (i.e., the third dimension) in pictures is likely to be problematic for inexperienced viewers, although even this difficulty is by no means inevitable.

Explaining the Research

What could account for these findings? As Hagen (1980) has pointed out, investigators working in this area of research tend to be concerned with empirical details rather than with theoretical explication or synthesis. However, even before most of this research had been conducted, Gombrich (1960) had articulated a theoretical framework which turns out, in retrospect, to have been a promising way of approaching the research findings. The basic hypothesis explored by Gombrich was that, even if a picture does not look very much like the thing it is meant to represent, it may nonetheless provide to the eye and brain informational cues which are similar to those used in the perception of the real world. For example, perhaps the structural information contained in a stick figure (the relative length and position of the various body parts) is meaningful by itself to any perceiver (including someone who has never seen such a picture before). In other words, a crucial assumption underlying this approach is that of *modularity*: Perhaps our perceptual apparatus operates on the basis of discrete informational cues (about structure, about depth, etc.), so that, even if some of these cues are absent from a particular visual display (e.g., a picture), perception can proceed unimpeded with regard to the other parts of the display.

This view of picture perception is very different from the kind of approach exemplified by Goodman and his followers. It suggests that a broad range of pictorial conventions which may seem at first blush to be arbitrary are actually derivatives of real-world perceptual processes. It also suggests, therefore, that the continuity between skills of pictorial interpretation -- which may be seen as one of the building blocks of visual literacy -- and real-world perceptual processes may be much greater than someone like Goodman might imagine.

Although Gombrich's assumptions appeared to fit in well with a variety of observations in perceptual psychology as well as art history, he himself was initially quite tentative about these assumptions, because the state of scientific knowledge about visual perception was too inconclusive at the time to allow him to take a firm position. His subsequent encounter with the work of J.J. Gibson (1982) has dispelled much of this tentativeness (e.g., see Gombrich, 1984). However, for a fuller assessment of the adequacy of Gombrich's approach, we must turn to more recent developments in perceptual psychology. The past decade has seen major advances in the theory of vision, stemming most notably from the work of the late David Marr of M.I.T. (See Marr, 1982, for the most accessible account of Marr's basic theory.) These theoretical advances have resulted in a significant reconceptualization of the visual process, with direct implications for the issue which concerns us here, viz., the degree of continuity between real-world (unmediated) perception and the perception of pictures.

The Visual Process

The specific aspect of vision theory from which this connection emerges has to do with the brain's "translation" of the retinal image into a mental representation of identifiable objects in three-dimensional space. Marr's model of this process comprises three discrete steps. In the first step, the brain scans the visual information transmitted to it from the retina -- i.e., essentially a two-dimensional array of color and light values -- in search of discontinuities corresponding to the outlines of objects or the edges of surfaces. The end result of this stage of the visual process is a mental representation which Marr labels the "primal sketch" and which can be thought of as being equivalent to an outline drawing of the scene which the eyes are looking at. All further mental operations aimed at identifying the objects in the scene and inferring the scene's three-dimensional properties are conducted on the basis of this outline representation.

In other words, here we have an affirmation of Gombrich's assumption of modularity in visual processing: The ability to recognize objects and see three-dimensionally without reference to most details of light or color is actually built into our real-world visual apparatus. On the basis of this principle it seems reasonable to hypothesize that a pictorially inexperienced viewer should not in fact be hampered by the absence of color and/or shading in black-and-white photographs, outline drawings, etc., and, as already noted, systematic research -- as opposed to anecdotal evidence -- has supported this hypothesis (e.g., see Dusenbury, 1990; Hochberg and Brooks, 1962; Spain, 1983; Zimmerman and Hochberg, 1970).

Depth Perception

Step two in Marr's model of the visual process is concerned with computing depth -- i.e., the third dimension -- in the outline representation resulting from step one. This computation makes use of several types of information, or "depth cues," but for present purposes the following four are the most relevant ones:
(1) *Binocular Disparity*. Because the retinal image registered in the left eye is slightly different from that of the right eye (since the two eyes' points of view differ), and because the

magnitude of this difference in retinal images is inversely related to our distance from whatever it is we are looking at, the brain can use this difference in its computation of depth.

(2) *Motion Parallax*. Any change in the relative position of the viewer vis-a-vis the thing being viewed is accompanied by a change in the retinal image of that thing. Because the magnitude of the retinal change is also inversely related to distance (between viewer and thing viewed), it too can serve as a depth cue.

(3) *Texture Gradients*. Whenever we look at a scene containing a regular texture or pattern (e.g., a tiled floor, a picket fence, railroad ties, or even such naturally-occurring patterns as uniform-sized pebbles on a beach), the retinal image of that pattern will increase in density as distance increases. These variations in apparent density can serve as informational cues regarding the distance between the observer and the various parts of the scene she or he is looking at.

(4) *Occlusion*. When one object blocks our view of another object, we know that the former is closer to us than the latter. This is a trivially obvious, but nonetheless compelling, source of information about the three-dimensional properties of the scene we are looking at.

It should be readily apparent that the first two of these depth cues cannot possibly be incorporated into any ordinary, non-holographic still picture (although pairs of images viewed through 3-D glasses or other stereoscopic devices can of course mimic the effect of binocular disparity.) If our brain required the simultaneous operation of all four cues in order to compute depth, we would be forced to conclude that the perception of depth in still images must be an impossible task for pictorially inexperienced viewers. However, once again the principle of modularity comes into play here.

As Marr and others have noted (e.g., see Kubovy, 1986), there is good reason to believe that these depth cues can indeed operate independently of one another. Consequently, if a particular picture contains enough information about depth in the form of texture gradients and/or occlusion, inexperienced viewers should *not* find it difficult to see depth in that picture. A picture meeting these criteria has been tested in a study conducted in New Guinea by Cook (1981), and the results supported this assumption. On the other hand, of course, when texture gradients and occlusion are either entirely absent from a picture or are not sufficiently informative to generate an adequate sense of three-dimensional space, we should not be surprised to find that inexperienced viewers do not perceive depth in that picture. This assumption, too, has been confirmed experimentally (see Hudson, 1967; also Hagen and Johnson, 1977; Hamdi et al., 1982; Mshelia and Lapidus, 1990).

Object Identification

The third and final step in the visual process which concerns us here entails the identification of the various objects in the visual field. This part of the process is not yet understood as well as the previous two, but what is known about it does permit us to say that Gombrich appears to have been correct in speculating that structure alone -- without surface details -- might be a sufficient basis for object recognition.

As described by Marr, the brain's activity at this stage in the visual process consists of "reducing" the outlines of objects into a more elementary representational form in which only the basic underlying structure is retained, while many incidental details are discarded. (In other words, at this point the brain's processing of visual information goes beyond the outline format which was the result of step one.) These more basic structural representations are then matched against a "dictionary" of object structures in the brain's memory. Thus, two different outlines will lead to the same ultimate labelling of an object so long as their underlying structures match; and it is this feature of real-world vision which is assumed to account for our ability to recognize a particular object in any number of different circumstances (e.g., different points of view, different levels of illumination, etc.).

When it comes to specifically *pictorial* variations in the appearance of an object, then -- e.g., the transformations entailed in stick figures, cartoons, sketches, etc. -- Marr's theory suggests that a viewer's ability to interpret such pictures may be an extension of an everyday, real-world perceptual skill, rather than something one has to learn through previous experience with similar pictures. It follows, therefore, that inexperienced viewers should be able to recognize the objects in such pictures without much difficulty -- and, as noted earlier, this has indeed been found to be true in studies of this issue (Cook, 1981; Kennedy and Ross, 1975).

Implications

We have seen that a broad variety of pictorial styles which may seem, intuitively, to entail major departures from the appearance of the real world -- e.g., black-and-white photographs, outline drawings, stick figures, cartoons -- do not appear to pose significant interpretational obstacles to viewers unfamiliar with those styles. This observation, which has been confirmed by a growing body of empirical investigations, goes against the implications of certain oft-cited anecdotal accounts about first-time viewers' responses to photographs and other kinds of images. Does the argument developed in this paper imply that these anecdotal accounts were exaggerations or even fabrications? Not necessarily. As Deregowski et al. (1972) have shown in a detailed study of this issue, it is quite likely that the viewers' puzzlement reported in these accounts was due to lack of familiarity with the pictorial media -- typically, paper -- rather than being a response to the actual content of the pictures. This view has been supported by recent research in an isolated rural area in Kenya (R. Hobbs, personal communication, 1992).

What are the broader implications of the issues examined in this paper? It has been argued that the skills required for the interpretation of the conventions of pictorial representation may be, to a significant extent, derivatives of real-world perceptual processes, rather than arbitrary, picture-specific mental habits. To the extent that this is true, pictures may be thought of as being, in that regard, a more readily accessible mode of communication than language, whose semantic and syntactic codes are almost exclusively arbitrary. Does this greater accessibility of pictorial communication imply that the social need for visual literacy is less than the need for verbal skills? If anything, the contrary should be true. Precisely because of their greater accessibility to the untutored viewer, images as a mode of communication may be unmatched in their capacity for manipulation and misinformation (see Messaris, 1992; in press). The need for visual literacy in its broadest sense -- i.e., as reflective, critical awareness of visual conventions and their uses -- should be correspondingly acute.

References

Cook, B. (1981). *Understanding pictures in Papua New Guinea*. Elgin, IL: David C. Cook Foundation.

Deregowski, J.B. (1980). *Illusions, patterns, and pictures: A cross-cultural perspective*. New York: Academic Press.

Deregowski, J.B.; Muldrow, E.S.; and Muldrow, W.F. (1972). Pictorial recognition in a remote Ethiopian population. *Perception*, 1: 417-425.

Dusenbury, K.A. (1990). *Inexperienced viewers' understanding of pictorial materials: The results of a survey in Lesotho, Africa*. M.A. Thesis, Annenberg School for Communication, University of Pennsylvania.

Gibson, J.J. (1982). *Reasons for realism: Selected essays of James J. Gibson*. Ed. by E. Reed and R. Jones. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gombrich, E.H. (1960). *Art and illusion: A study in the psychology of pictorial representation*. Princeton, NJ: Princeton University Press.

Gombrich, E.H. (1984). Representation and misrepresentation. *Critical Inquiry*, 11: 195-201.

Goodman, N. (1976). *Languages of art: An approach to a theory of symbols*. 2nd ed. Indianapolis: Hackett.

Hagen, M.A. (1980). Generative theory: A perceptual theory of pictorial representation. In M.A. Hagen (Ed.), *The perception of pictures*, Vol. II. New York: Academic Press. 3-46.

Hagen, M.A., and Johnson, M.M. (1977). Hudson pictorial depth perception test: Cultural content and question with a Western sample. *Journal of Social Psychology*, 101: 3-11.

Hamdi, N.; Knirk, F.; and Michael, W.B. (1982). Differences between American and Arabic children in performance on measures of pictorial depth perception: Implications for valid interpretation of test scores based on items reflecting dissimilar cultural content. *Educational and Psychological Measurement*, 42: 285-296.

Hochberg, J., and Brooks, V. (1962). Pictorial recognition as an unlearned ability: A study of one child's performance. *American Journal of Psychology*, 75: 624-628.

Hudson, W. (1967). The study of the problem of pictorial perception among unacculturated groups. *International Journal of Psychology*, 2: 89-107.

Kennedy, J.M., and Ross, A.S. (1975). Outline picture perception by the Songe of Papua. *Perception*, 4: 391-406.

Kubovy, M. (1986). *The psychology of perspective and Renaissance art*. New York: Cambridge University Press.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: W.H. Freeman.

Messaris, P. (1992). Visual "manipulation": Visual means of affecting responses to images. *Communication*, 13: 181-195.

Messaris, P. (In press). *Visual Literacy: Image, Mind, and Reality*. Boulder, CO: Westview Press.

Mshelia, A.Y., and Lapidus, L.B. (1990). Depth picture perception in relation to cognitive style and training in non-Western children. *Journal of Cross-Cultural Psychology*, 21(4): 424-433.

Panofsky, E. (1925). *Perspective as symbolic form*. Tr. by C.S. Wood. New York: Zone Books, 1991.

Segall, M.H.; Campbell, D.T.; and Herskovits, M.J. (1966). *The influence of culture on visual perception*. Indianapolis: Bobbs-Merrill.

Snyder, J. (1980). Picturing vision. *Critical Inquiry*, 6(3): 499-526.

Spain, S. (1983). *Factors affecting pictorial comprehension in non-literates: Results of a survey in the Gambia, West Africa*. M.A. thesis, Annenberg School for Communication, University of Pennsylvania.

Wartofsky, M. (1984). The paradox of painting: Pictorial representation and the dimensionality of visual space. *Social Research*, 51(4): 863-883.

Zimmerman, R., and Hochberg, J. (1970). Responses of infant monkeys to pictorial representations of a learned visual discrimination. *Psychonomic Science*, 18(5): 307-308.